Symbolic Inference for Non-Horn Knowledge Bases With Fuzzy Predicates

Alexander Sakharov

Polynomial Computer Algebra

May 2-7, 2022

Non-Horn Knowledge Bases

- Atoms: $P(t_1,\ldots,t_n)$ (P - predicate, $t_1,\ldots,t_n$ – terms)
- Literals are atoms or their negations
- Non-Horn Knowledge Bases (KB):

  Facts are literals

  Rules have the form: $A \Leftarrow A_1 \wedge \ldots \wedge A_n$ ($A, A_1, \ldots, A_n$ – literals)

  Inference goals are ground literals
- Principle of double negation: $\neg\neg A \equiv A$
- First-order logic (FOL) axioms and inference rules
- Non-Horn KBs: more powerful than Horn KBs and non-controversial in contrast to normal logic programs

Fuzzy Knowledge Bases

- Fuzzy KBs: Predicates are fuzzy. Their truth values are real numbers from interval *[-1,1]* (*-1* – false, *1* – true)
- *|A|* denotes the truth value of *A*
- Some predicates are implemented by possibly partial programs or by neural networks. The truth values of their ground atoms (i.e. without variables) are calculated. All other predicates are derivable.
- KB facts have truth values higher than a certain threshold $h > 0$
- Inference from fuzzy KBs is based on fuzzy generalized Modus Ponens and is limited to forward chaining
- The goal of inference is either to estimate the truth values of literals or to determine fuzzy-set membership (the latter additionally involves fuzzification and defuzzification)

Fuzzy Logic

- Traditional negation truth function for fuzzy KBs: $|\neg A| = - |A|$
- Traditional conjunction truth function for fuzzy KBs (Godel t-norm):

  $$| A_1 \wedge \ldots \wedge A_n | = min\{|A_1|,\ldots,|A_n|\}$$
- Other truth functions: other t-norms for conjunction (product, Lukasiewicz) and truth functions for other connectives based on t-norms
- Fuzzy connectives may not be dependent
- Non-Horn KBs: The use of the negation truth function is limited to the calculation of truth values of negative literals. The use of the conjunction truth function is limited to the calculation of truth values of KB rule bodies. Truth functions for other connectives are not necessarily used.
- Semantics of KB rules: the truth value of the rule body is a lower bound of the truth value of the head. (This is similar to fuzzy generalized Modus Ponens.)

Reductio Ad Absurdum (RAA)

- The principle of RAA (aka reasoning by contradiction) states that *A* is derivable if it is derivable from the hypothesis $\neg A$
- In the presence of fuzzy predicates, it seems faulty to apply RAA, i.e. reason by contradiction, when the truth value of the hypothesis is close to zero:

Here is reasoning by contradiction in two-valued logic from the two KB rules $P \Leftarrow Q$ and $P \Leftarrow \neg Q$:

Suppose *P* is false. The first rule implies that *Q* is false, and hence *P* is true by the second rule.

Now consider fuzzy logic and suppose $|P| = 0$. If $|Q| = 0$ as well, then both rules are satisfied, but they do not provide any evidence that *P* is true or $|P| > 0$ at least.

Our Goal and Method

- The goal is to obtain lower bounds of the truth values of ground literals given that these truth values satisfy the semantics of KB rules

    This goal is achieved by the following:

- Ground literals are derived in a calculus that characterizes inference from the non-Horn KB without RAA
- Symbolic expressions (terms) are built from derivations of ground literals in this calculus
- Lower bounds of the truth values of ground literals of fuzzy predicates are calculated by evaluating these symbolic expressions

Sequent Calculi

Antecedents are multisets of formulas, succedents are formulas

Structural rule:

$$\frac{\Gamma \vdash A \qquad A, \Pi \vdash B}{\Gamma, \Pi \vdash B}\ \textit{cut}$$

Logical rule:

$$\frac{A, \Gamma \vdash B}{\sim B, \Gamma \vdash \sim A}\ \textit{swap}$$

Non-logical axioms:

KB rules: $A_1 ,\ldots, A_n \vdash A$

KB facts: $\vdash A$

Variables can be replaced by terms in instances of these axioms

The conclusions of swap applied to KB rules are called contrapositives

Sequent Calculi (continued)

Definition. *L_{cs} is the set of sequent calculi in which formulas are literals, the structural rule is cut, the logical rule is swap whose premises are axioms, non-logical axioms represent KB facts and rules, no logical axioms.*

Theorem 1. *L_{cs} is sound and complete with respect to the derivation of ground literals in FOL without RAA.*

It is known that ground literal *L* is derivable from KB facts and rules in FOL without RAA if and only if *~L* is refutable by resolution in which the factoring rule is not used and at least one premise of every resolution step is not *~L* or its descendant.

Resolution -> L_{cs}: ground the resolution refutation; exclude the step that resolves *~L*; turn resolution steps to cut; insert swap

L_{cs} -> Resolution: due to the lifting lemma, any L_{cs} derivation can be transformed to a resolution derivation with such endclause *L'* that *L* is its instance; turn cut to resolution steps; add a step resolving *L'* and *~L*

Rule and Contrapositive Truth Functions

Due to the double negation principle for non-Horn KBs, the negation truth function should be an involution. This does not leave other good choices beyond $n(x) = -x$.

The functions $[-1,1]^k \rightarrow [-1,1]$ whose arguments are the truth values of literals in the bodies of KB rules or contrapositives will be called rule or contrapositive truth functions, respectively.

For the Godel t-norm, the semantics of rule $A_0 \Leftarrow A_1 \wedge \ldots \wedge A_k$ implies that $|{\sim}A_j| \geq |{\sim}A_0|$ provided that $|A_0|$ is negative and the other $|A_i|$ are positive.

$g(x_0, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k) = x_0$ – the contrapositive truth function

Other truth functions than the Godel t-norm may be more appropriate for the bodies of non-Horn rules with fuzzy predicates. The product and Lukasiewicz t-norms do not seem a good choice.

An individual truth function can be provided for any KB rule.

Truth Functions (continued)

The following function seems reasonable:

$s(x_1,\ldots,x_k) = \sqrt[k]{(x_1+1)\ldots(xk+1)}\ -1$

($x_1,\ldots,x_k$ are the truth values of the body literals)

The semantics of rule $A_0 \Leftarrow A_1 \wedge \ldots \wedge A_k$: $|A_0| \geq s(|A_1|,\ldots,|A_k|)$

This inequality can be symbolically solved for $|A_j|$ provided that $|A_0|$ is negative and the other $|A_i|$ are positive. The right-hand side of the solution $|{\sim}A_j| \geq \ldots$ gives the contrapositive truth function.

Definition. *Rule or contrapositive truth function s is called proper if $s(1, \ldots, 1) = 1$, $s(0, \ldots, 0) = 0$, and for $j = 1\ldots k$, $h \leq s(x_1, \ldots, x_j, \ldots, x_k) \leq s(x_1, \ldots, x'_j, \ldots, x_k)$ given that $x_j \leq x'_j$ and $x_1 \geq h, \ldots, x_k \geq h$.*

The rule and contrapositive functions specified earlier satisfy the conditions of this definition

Truth Value Approximation

Term $n(\tau)$ is recursively defined for all ground derivations τ in L_{cs}:

- If τ is ground instance A of a KB fact, then $n(\tau) = |A|$ ($|A|$ is a constant).
- If τ is ground instance $A_0 \Leftarrow A_1, \ldots, A_k$ of KB rule and s is the truth function for this KB rule, then $n(\tau) = s(a_1, \ldots, a_k)$.
- If the last rule of τ is swap with the conclusion $A_0, A_1, \ldots, A_{j-1}, A_{j+1}, \ldots, A_k \vdash A_j$ and s' is the truth function for the respective contrapositive, then $n(\tau) = s'(a_0, a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_k)$.
- If the last rule of τ is cut with premises $A_1, \ldots, A_k \vdash E$ and $E, C_1, \ldots, C_m \vdash D$, then $n(\tau) = n(\nu)\{e \rightarrow n(\mu)\}$. Here, μ and ν are the parts of τ whose endsequents are the premises of this cut

Truth Value Approximation (continued)

Theorem 2. *If τ is a ground L_{cs} derivation of literal G and all rule and contrapositive functions are proper, then $|G| \geq n(\tau) \geq h$.*

The following three statements are proved by induction on the depth of derivations:

- $n(\tau)$ does not contain variables for any derivation τ with the endsequent $\vdash G$.
- $n(\tau)$ is an increasing function with respect to every variable in it.
- If $A_1, \ldots, A_k \vdash D$ is the endsequent of derivation μ, then $|D| \geq n(\mu)\{a_1 \rightarrow |A_1|, \ldots, a_k \rightarrow |A_k|\}$. As a corollary, $|G| \geq n(\tau) \geq h$.

<u>Our goal is achieved: *n(τ)* gives a lower bound</u>

Time Complexity

- The proof of Theorem 1 shows that resolution refutations without factoring can be transformed to L_{cs} derivations in a single preorder traversal of the resolution refutations. Therefore, the time complexity of this transformation is linear in the size of the derivations.
- We assume that the time complexity of algorithms implementing rule and contrapositive functions is linear in the number of function arguments. The size of $n(\tau)$ is proportional to the size of τ . Hence, the calculation of a lower bound of $|G|$ takes a linear time of the size of G's derivation in L_{cs}.

Conclusion

- A simple proof-theoretic characterization of non-Horn KBs with fuzzy predicates is presented
- Our method of obtaining lower bounds of the truth values of ground literals of fuzzy predicates is applicable to a variety of truth functions for fuzzy KB rules
- Inference without RAA is more powerful than forward chaining that is normally used for fuzzy KBs
- Efficient inference methods such as ordered resolution can be directly utilized for deriving ground literals, and the remaining computations of the truth value bounds take a linear time of the size of resolution refutations

Q & A